PERGAMON

Vision Research 42 (2002) 2499–2508

Vision Research

www.elsevier.com/locate/visres

# On temporal hyperacuity in the human visual system

J.M. Zanker [a,*], J.P. Harris [b]

[a] Department of Psychology, University of London, Royal Holloway, Egham, Surrey TW20 0EX, UK
[b] Department of Psychology, University of Reading, Earley Gate, Reading RG6 6AL, UK

Received 29 January 2002; received in revised form 13 May 2002

**Abstract**

The spatial grain of the human visual system has always been a central topic for visual sciences, and the optical and physiological basis of perceptual limitations are well described. In particular, we have thorough accounts of spatial hyperacuity, which refers to a precision in the spatial localisation of stimulus contours that is better than the photoreceptor grain that determines spatial resolution. However, although the temporal resolution of the human visual system is comparably well described, we have almost no direct knowledge about the precision of localising visual stimuli in time in the absence of correlated spatial cues. The present study addresses this question by comparing directly the temporal resolution of human observers with their temporal acuity as measured in a temporal bisection task. Despite some improvement with practice, temporal acuity in this task does not fall below 20–30 ms in the best case, which is similar to the temporal resolution limit, and performance does not improve for comparison tasks with multiple stimulus presentations. The absence of visual hyperacuity for purely temporal modulations as tested here contrasts with processing limitations for other types of visual information in comparable tasks, and with other sensory modalities, in particular to those of the auditory system. Such differences can be interpreted in the context of the ecological requirements for organising behaviour, and the functional design of nervous systems.

*Keywords:* Vision; Spatial grain; Temporal resolution; Temporal accuracy; Hyperacuity; Motion

## 1. Introduction

Two classical measures are used to describe the spatial performance limits of the human visual system. (a) Grating *resolution* refers to the finest periodic stimulus that can be perceived. A similar task is that of two-point resolution, which tests whether two adjacent targets, such as spots of light, can be separated from each other. (b) *Positional acuity* is typically studied in a Vernier task in which the alignment of two short lines has to be judged or in a bisection task in which a target has to be positioned exactly in the middle of two other stimulus elements. It is well known that the position of a visual target can be estimated with a much higher precision than the smallest detectable gap suggested by spatial resolution (Westheimer, 1984). Since this performance also goes beyond the anatomical grain of the human visual system it is called hyperacuity and is attributed to cortical processing of neighbourhood relationships in retinotopic maps. Hyperacuity can be found for a variety of stimulus features in spatial vision, such as luminance, colour, stereo, but also for targets defined by motion or local contrast. Thus the mechanisms underlying spatial hyperacuity can be interpreted as a rather general type of processing using various sources of low-level sensory information (Morgan, 1991).

Is a similar dichotomy between resolution and positional acuity useful in describing the performance limits of temporal vision? In other words, is there temporal hyperacuity? Resolution limits for periodic modulation of light intensity are, again, well described, and the flicker-fusion frequency for stabilised images lies between 30 and 60 Hz, depending on the exact experimental conditions (Kelly, 1972). It is generally accepted that flicker fusion relates to the dynamic properties of photoreceptors and gives a reasonable estimate of the temporal limit of vision, in that the human visual system is not expected to process faster changes in visual stimuli. However, it would be

---

[*] Corresponding author.
*E-mail addresses:* j.zanker@rhbnc.ac.uk (J.M. Zanker), j.p.harris@reading.ac.uk (J.P. Harris).

0042-6989/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S0042-6989(02)00301-2

interesting to know whether location in time, for example the exact phase of periodic modulations, can be perceived with a precision which goes beyond this temporal resolution limit.

In the context of other visual phenomena, in particular in motion vision, the temporal grain of the visual system seems to be higher than the resolution limit of about 40 Hz, corresponding to a period length of 25 ms. In his classical study, Exner (1875) found that minimum time differences in temporal order discrimination go down to 44 ms if both targets are at the same position in space, but to 17 ms if they are separated and motion is perceived. The discrimination of two separate lights flashed in neighbouring spatial locations under optimum conditions (which include presentation in the light-adapted periphery) can be made with a precision of 2–5 ms, using motion cues (Sweet, 1953). Temporal displacement thresholds tested directly for apparent motion stimuli (i.e., the successive presentation of two lines in close spatial proximity), may reduce to 2–4 ms (Westheimer & McKee, 1977) or 3–5 ms, depending on the spatial separation of the lines (Wehrhahn & Rapf, 1992). Similar results are achieved for Vernier displacement thresholds in spatio-temporal interpolation, in which a sequence of vertical lines is presented in apparent motion at consecutive horizontal positions. When the lower half of the lines is presented with a tiny temporal delay—but at the same horizontal position as the upper half—a Vernier displacement is perceived. Thresholds for detecting such apparent displacements reduce to 10–20 arcsec under optimal conditions, which converts to a detection of about 1–2 ms delay (Burr, 1979; Fahle & Poggio, 1981; Morgan & Watt, 1983). Similar lower limits in the milliseconds range are achieved for inter-ocular phase shifts which convert to perceived depth modulation in a binocular spatio-temporal interpolation task (Morgan & Watt, 1982). It should be noted however that in this class of stimuli the extraction of depth information is not limited by the absolute delay between the input to the two eyes, but rather the temporal phase of the modulation function (Morgan & Fahle, 2000).

Furthermore, two-pulse sequences in which the targets are presented at the same position but are defined by two different colours can be perceived with high precision: their temporal order can be discriminated with a precision of 1–2 ms, if presented to a single eye. No motion is involved, but temporal patterns can be discriminated on the basis of the colour change (as a stronger sensation of one particular colour at the end of a sequence), and thus cannot be regarded as pure test of temporal acuity. This explanation of the very high precision achieved in that paradigm seems to be supported by the additional observation that sensitivity for offset-asynchrony is higher than for onset-asynchrony (Yund & Efron, 1974). Another indication of high temporal acuity related to colour vision arises from the flicker-induced colours of Benham's Top (Both & Campenhausen, 1978; Tritsch, 1992).

Most of these phenomena, however, involve some combination of spatial and temporal displacement, and thus are only indirect evidence for high temporal resolution, which may be mediated through some sort of spatio-temporal filters. The crucial problem with temporal order discriminations is that the two events somehow have to be identified by the observers, and therefore need to be labelled by some target property or kept separate along a certain stimulus dimension. Apart from the danger of generating local motion signals to which humans are extremely sensitive, the involvement of different stimulus attributes (e.g., colours) may give rise to variations in non-temporal cues such as processing gain or speed. Thus the question arises of how temporal acuity, and temporal hyperacuity, can be tested in a task that is free from other cues. We demonstrate below how this issue can be addressed with a simple experiment: observers evaluate the temporal position of one event with respect to two others in a temporal bisection task. Temporal bisection tasks have been used extensively in other contexts, but have typically involved the presentation of two intervals separated by a gap, and have been aimed at discovering the properties of the internal clock, not at measuring the absolute precision of bisection, or establishing its relationship to visual temporal resolution (e.g., Wearden & Bray, 2001).

## 2. Methods

### 2.1. Apparatus

A yellow LED was driven by a Mono-Flop, connected to the serial port of an IBM-compatible computer, and using its power supply. The LED was switched on and off by writing single ASCII characters with a BASIC program to the serial port, thus triggering the Mono-Flop. The device was developed by SMR W. Junger, Gomaringen, Germany. The shape of the pulses and their temporal precision was verified with an oscilloscope, confirming that rectangular pulses of 6–8 ms duration and square-wave modulations up to 70 Hz are reliably produced using an Intel 386 processor or higher, at a serial baud rate of 9600 bps. The luminance of the LED followed the rise and fall of the driving voltage with high precision, so that tiny imperfections were well beyond the resolution limit of the visual system. In this connection, it should be kept in mind that the inherent temporal filtering of the photoreceptors and ganglion cells will lead to a neural representation of the LED pulse which removes all high-frequency components from the input signal. Such temporal blur caused by the dynamic limitations of the sensory apparatus can be regarded as the cause of the temporal resolution limit, just as optical blur is the physical limiting factor for spatial resolution. Thus, in both domains, we are dealing with low-pass filtered signals, which is a crucial precondition for the operation of spatial hyperacuity mechanisms (Morgan, 1991).

### 2.2. Stimuli

In the main experiment, a sequence of three consecutive light pulses (8 ms duration) of identical intensity was presented. The second pulse varied in its temporal position, i.e. its delay from the first pulse, whereas the interval between the first and last pulse (overall time interval) was constant for a given experiment. The subject's task was to decide whether the second pulse was closer in time to the first or to the last pulse. This temporal bisection task provides the observer with no non-temporal cues that could be used to make the perceptual decision. In particular, no apparent motion cues or attribute change cues are introduced, because all three light flashes share the same position, colour and intensity. However it has to be kept in mind that single, short events like individual LED light pulses may be difficult to perceive as such, and some practice may be required before the observers achieve optimal performance.

For our first measurement of temporal resolution, two consecutive light pulses each of 8 ms duration, separated by a brief gap, were paired in random order with a single pulse that lasted twice as long as each of the two pulses (i.e., delivered the same light energy). The time gap between the two pulses was varied (usually in multiples of 4 ms), and the subjects had to decide in a two-interval forced choice (2IFC), whether they saw the double pulse in the first or second interval. Similarly in a second comparison experiment, observers discriminated in a 2IFC paradigm periodic light modulations (square-wave profiles generated by switching the LED on and off) of variable frequency from a quasi-continuous light of identical average intensity (in fact a modulation of 70 Hz). This test was used to estimate flicker fusion frequency under the same experimental conditions as in the temporal bisection task.

### 2.3. Procedures

Subjects were seated in a moderately illuminated room and fixated the LED binocularly through natural pupils from a distance of about 40 cm. In a constant stimulus procedure, the temporal variable (delay, gap width or frequency) was set to 11 different values. A single experimental block comprised 16 presentations of each stimulus in randomised order. The subject's decision, signalled via the keyboard, was stored by the computer program for later analysis. No feedback about the accuracy of judgements was given. The psychometric curves derived from these data sets were used to calculate temporal acuity thresholds by means of Probit analysis (Finney, 1962), usually for the average of four experimental blocks. The two authors and three observers, who were experienced psychophysical observers but naïve to the purpose of the experiments, took part in this study (three male, two female; age ranging between 23 and 55 years). Their visual acuity was normal or was corrected optically to normal.

## 3. Results

Each of the main figures shows the complete data set, i.e. the full psychometric curves, from a single subject for a typical stimulus condition, and thresholds calculated from such psychometric curves for four subjects and a variety of stimulus conditions. The fifth subject turned out to be unable to solve the temporal bisection task (while achieving 'normal' temporal resolution), generating largely unbalanced decisions even after substantial practice with this task, and therefore could not be included in quantitative analysis. Each experiment was usually repeated in sets of four separate blocks (data indicated by small symbols and thin lines in the psychometric curves) so as to provide 64 decisions for each condition for any threshold estimate. The three sections below compare directly the temporal processing limits for the three tasks, in which the same equipment was used under identical conditions.

### 3.1. Temporal bisection

Three LED pulses were presented (see inset of Fig. 1a) with an overall time interval of 512 or 1024 ms (additional configurations were tested with subject JMZ). The temporal distance of the second flash relative to the centre between the two others was varied in multiples of 16 ms. The subject had to decide in a 2IFC paradigm whether the centre pulse was perceived 'early' (i.e., closer to the first flash) or 'late' (i.e., closer to the last flash). The average psychometric function shown for subject JMZ in Fig. 1a shows reliable decisions for about 3/4 of the temporal positions of the second flash, and a transition from 'front' to 'back' over a range of less than 128 ms. For the centre position (i.e., same distance to first and last flash) decisions are virtually balanced in this example. The points on the psychometric function at which 75% and 25% of the decisions that the second pulse was 'early' are at a temporal distance from the centre of 28.9 and 21.7 ms, respectively. From the average of these two values, temporal position acuity can be estimated as the 'just noticeable difference' (JND), being 25.3 ms in this example. The point of subjective equality (PSE) of the psychometric curve indicates any systematic deviation from the centre position in the

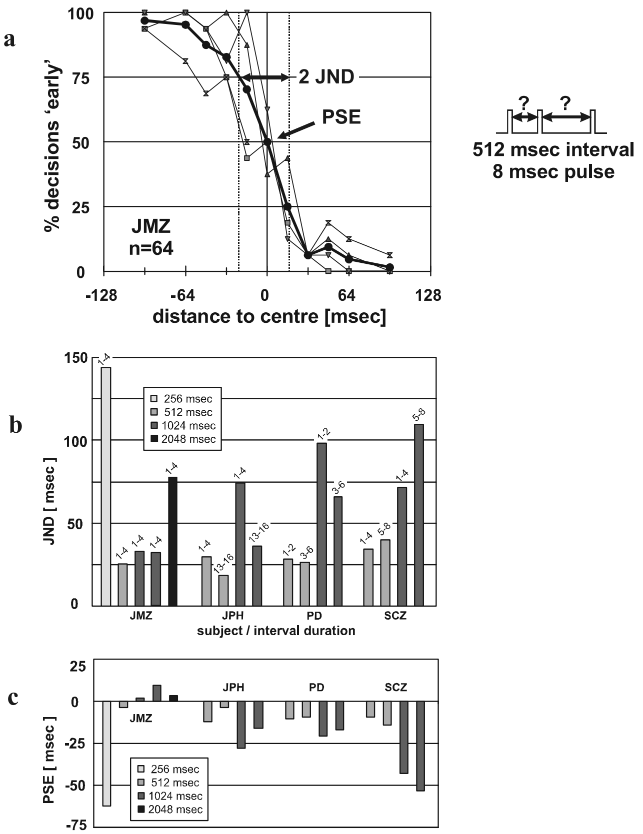

Fig. 1. (a) Psychometric function for subject JMZ showing the frequency of decisions that the second of three light pulses was closer to the first than to the third pulse in the temporal bisection task (sketched as inset) as a function of the physical presentation time of the second pulse (given as distance from the centre). Small symbols and thin lines show the result from a single block; the black dots and thick lines show the average data from four blocks, which were used to estimate the PSE and JND. (b) JNDs indicating temporal position acuity, estimated for four subjects from such psychometric curves for various overall durations of the interval between the first and third light pulse (indicated by shading of the bars), and for several consecutive blocks (labels above the respective bars) in each condition. (c) PSEs indicating the systematic misjudgements of temporal position, estimated for the same conditions as in (b).

temporal bisection task. Negative values mean that the first, more distant, interval is perceived longer than the second, thus leading to an early bisection, whereas positive values indicate that the second, more recent, interval is perceived as longer by the observer. In the example shown in Fig. 1a the PSE is −3.6 ms, suggesting that this subject is not experiencing any major perceptual distortion of the time axis. The JND and PSE values are shown in Fig. 1b and c, respectively, for all four subjects and a variety of experimental conditions.

For longer and shorter overall intervals (256, 1024, 2048 ms) the task gets more difficult than for the optimum overall interval of 512 ms. The psychometric functions become flatter, temporal position acuity is correspondingly worse (about 144, 32, 78 ms for the highly trained subject JMZ, see Fig. 1b), and larger deviations of the PSE from the centre of the stimulus interval occur (see Fig. 1c). This observation corresponds to the subjective impression of the observers that shorter durations lead to difficulties with separating the first and second interval of each stimulus because everything is happening too fast, and that longer durations lead to a problem with memorising the duration of the two intervals, the first one being further away in time than the second one.

At the overall intervals that yield the smallest JNDs, 512 and 1024 ms, there is a training effect that can be observed for most subjects as a reduction of the average JNDs (usually calculated for four consecutive test blocks) when the experiment is repeated several times. In Fig. 2 this is shown in more detail for subject JPH, with JNDs and PSEs calculated separately for each block and plotted as a function of block number. JNDs gradually reduce (i.e., temporal acuity increases) through the course of 16 blocks, in particular for the less optimal stimulus with an overall interval of 1024 ms, with the largest changes in the first four blocks. This trend is accompanied by a tendency for the PSE (i.e., the

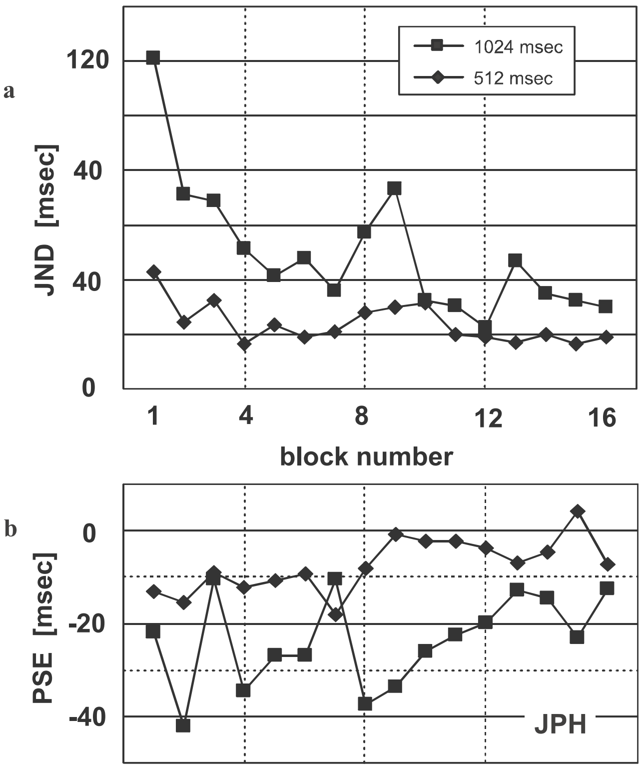

Fig. 2. JNDs (a) and PSEs (b) estimated for individual blocks for subject JPH who repeated the temporal bisection experiment 16 times to monitor practice effects (each data point is based on 16 decisions for each of 11 temporal positions). The two different symbols in each graph represent two overall durations of the bisection stimulus of 512 and 1024 ms, respectively. An initial reduction in systematic deviation from the actual centre (PSE) is accompanied by an improvement of temporal accuracy (decrease of JND), but after about four blocks further improvements are marginal.

systematic deviations) to reduce in a similar pattern. The same practice effects can be seen for one of the naïve observers (see Fig. 1b,c: PD), whereas the performance of the other deteriorates during the experiment (see Fig. 1b,c: SCZ), presumably due to a loss of motivation.

The general pattern that emerges from this set of results is apparent. Even under the most favourable experimental conditions (optimum duration of overall interval, extended periods of practice) the temporal acuity reached in the bisection task, which is defined as detecting reliably the deviation from a temporal centre position of the second light pulse in a sequence of three is in the range of 20–30 ms (Fig. 1b). Furthermore, our observers, if they were not highly overtrained or almost incapable of solving the experimental task, exhibited a slight tendency to perceive the first interval as longer than the second (Fig. 1c), which could be interpreted as a perceptual compression of time, perhaps via some effect of attention or arousal on the internal clock, as suggested by Rose and Summers (1995), who report an extensive study of the phenomenon. We investigated in two control experiments how the lower temporal position acuity limit of 20–30 ms relates to the temporal resolution of the visual system under the same experimental conditions.

### 3.2. Two-pulse resolution

In the first comparison experiment, the subject had to detect in a 2IFC paradigm two separate pulses in close succession. The average psychometric curve for subject JMZ shown in Fig. 3a is characterised by a smooth transition from random choice (around 50% correct) for very brief gaps between the two pulses (about 10 ms) to very reliable discrimination (close to 100% correct) for gap widths of 50–60 ms. The discrimination threshold, corresponding 75% correct decisions, was estimated by Probit analysis to be 27.1 ms under these particular experimental conditions.

The two-pulse resolution was measured for observer JMZ for two different ranges of gap widths (varied in steps of 8 ms between 8 and 120 ms, and in steps of 2 ms between 2 and 30 ms), leading to very similar estimates of temporal separation acuity (see Fig. 3b). The other subjects tested in this experiment show a similar range of thresholds, broadly scattered around 20–40 ms, with no substantial effects of practice. The large scatter in the data of Fig. 3a and the wide range of thresholds documented in Fig. 3b suggest that temporal two-point resolution is not a simple task under the present experimental conditions. A second, and perhaps more common, task was used to obtain a more reliable estimate of temporal resolution, as described in Section 3.3.

It should be noted that the overall duration from the first light onset to the last light offset varies together with the duration of the interval between the two pulses, and subjects may use the overall duration rather than the dark interval to discriminate the double pulse from the single pulse of identical overall energy. This, however, does not completely invalidate a two-pulse estimate of temporal resolution. Although detecting a change in the overall duration of an event may involve different mechanisms from detecting a brief blackout, each of the two tasks would provide an estimate for the precision of temporal measurement. In an attempt to maximise performance, the subject would employ whatever mechanism was more sensitive, thus reaching the lower limit of temporal encoding, i.e. the maximum sampling rate in time. If this ambiguity as to the underlying mechanisms may appear not completely satisfactory, one should remember that this test was included in the present experiments as an exact analogy to those commonly used in the spatial domain, where two-point resolution is a conventional measure of acuity. The same argument holds in that case—the task could be solved purely on the basis of the spatial spread of intensity, rather than by detecting a gap between two points—but measuring the spread requires high sampling density. In any case, this argument suggests the need for a second estimate of temporal resolution, which is described in the following section.

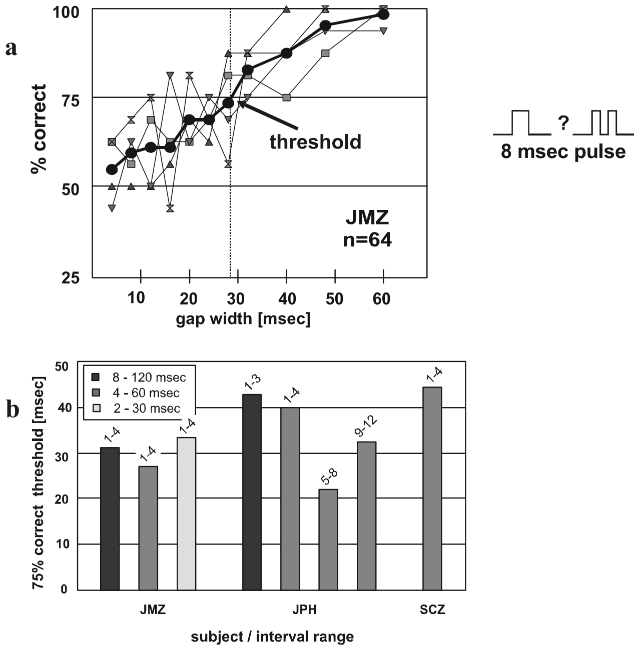

Fig. 3. (a) Psychometric function for subject JMZ, showing the frequency of correct detection of two brief light pulses in a 2IFC task (sketched as inset) as a function of the gap width between the pulses. (b) Detection thresholds estimated for three subjects from such psychometric curves for various ranges of gap widths (indicated by shading of the bars), and for several consecutive blocks in each condition. Two-pulse resolution in these experimental conditions is in the range of 20–40 ms. Conventions as in Fig. 1.

### *3.3. Temporal resolution*

The task in the second comparison experiment was to detect in a 2IFC paradigm the continuous square-wave modulation of the LED—a classical procedure to estimate the temporal resolution of the visual system. The average psychometric function, shown for subject JMZ in Fig. 4a, demonstrates, that under the present conditions the subject reaches a threshold of 75% correct decisions at a stimulus frequency of about 42 Hz. This corresponds to a period of less than 25 ms, meaning that the repetitively presented gaps are detected down to a value of little more than 10 ms. This performance clearly is better than the temporal position acuity observed in the temporal bisection task.

The resolution limits for all four subjects and various stimulus durations (500, 1000, and 2000 ms) are shown in Fig. 4b. It is clear that all thresholds are consistently in the range of 20–30 ms, no matter how long each stimulus was presented, and how well trained a subject was. These resolution limits compare favourably to commonly accepted flicker fusion frequencies in the central visual field (Kelly, 1972), and can be regarded as a conservative upper estimate of the temporal resolution limit, because it might be even higher under better experimental conditions.

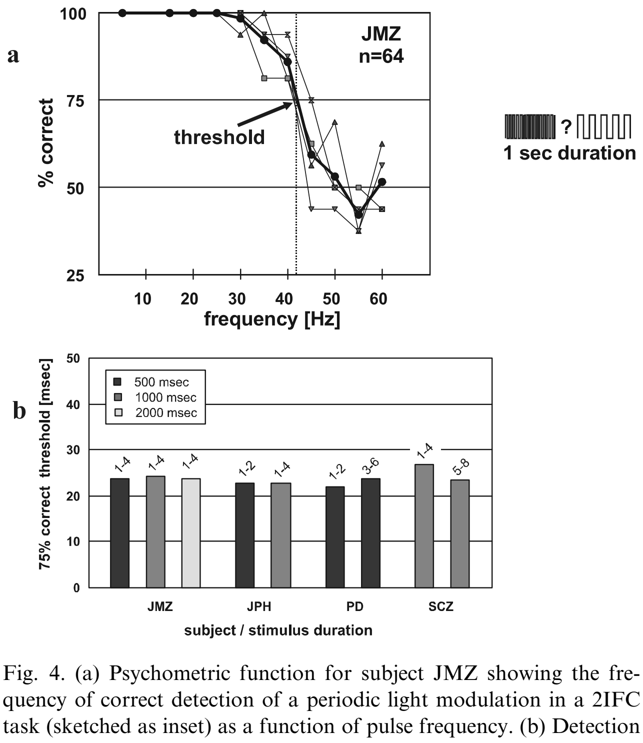

Fig. 4. (a) Psychometric function for subject JMZ showing the frequency of correct detection of a periodic light modulation in a 2IFC task (sketched as inset) as a function of pulse frequency. (b) Detection thresholds (expressed in ms as the period of the fundamental light modulation) estimated for four subjects from such psychometric curves for various overall stimulus durations (indicated by shading of the bars), and for several consecutive blocks in each condition. Temporal resolution measured in this classical task under the present experimental conditions is consistently in the range of 20–30 ms. Conventions as in Fig. 1.

### *3.4. Multiple bisection and detection of regularity*

At this stage of the work, it seemed that the thresholds in our temporal bisection task were rather similar to the resolution limit in our tasks with comparable stimuli, but it might be that performance would improve under more favourable conditions. This issue was addressed by two additional experiments, each carried out by one of the authors and a second, naïve subject.

A first concern is related to the fact that observers have to judge temporal location in the bisection task from a single event, whereas temporal resolution estimated from flicker fusion would allow for temporal integration across several events (i.e., intensity cycles), due to the nature of the repetitive stimulus. It should be kept in mind that the rationale behind the choice of stimulus conditions was comparability with the spatial domain, where the visibility of a displacement in a single Vernier bar is usually compared to the resolution of a repetitive grating, so that the comparison is between sensitivity for displacement at a single location and sensitivity to intensity modulation across a number of grating cycles. Although in the temporal domain we did not find differences in performance for analogues of two configurations that lead to substantial effects in the spatial domain, we wanted to know whether repetitive stimulation would lead to some improvement of temporal position thresholds.

In this comparison experiment, subjects had to solve a bisection task as before, but the two stimulus intervals (short—S or long—L) between light pulses are repeated four times, leading to a sequence SLSLSLSL or LSLSLSLS. The basic stimulus duration (S + L) was chosen to be 512 ms (which gave the best temporal position acuity for the single bisection judgements), and a variety of temporal positions of the centre pulse (S–L) was tested in a constant stimulus procedure as before. The subject had to decide whether the sequence started with a short interval and ended with a long one or vice versa. Such a repetitive temporal stimulus can be regarded as an analogue of a Vernier grating stimulus in the spatial domain, since both allow for a certain amount of integration, in one case over time, in the other over space. In fact, this experimental task turned out to be rather difficult for the observers, who found it hard to preserve the order of the various intervals, and to use them all in making a judgement. Contrary to the idea of temporal integration, the best strategy was to make a decision based on the first pair of intervals (i.e., stimulus cycle), and to largely ignore the three consecutive stimulus cycles, which acted almost like a mask. The psychometric curves from one subject in four repetitions of this experiment are shown in Fig. 5a, which

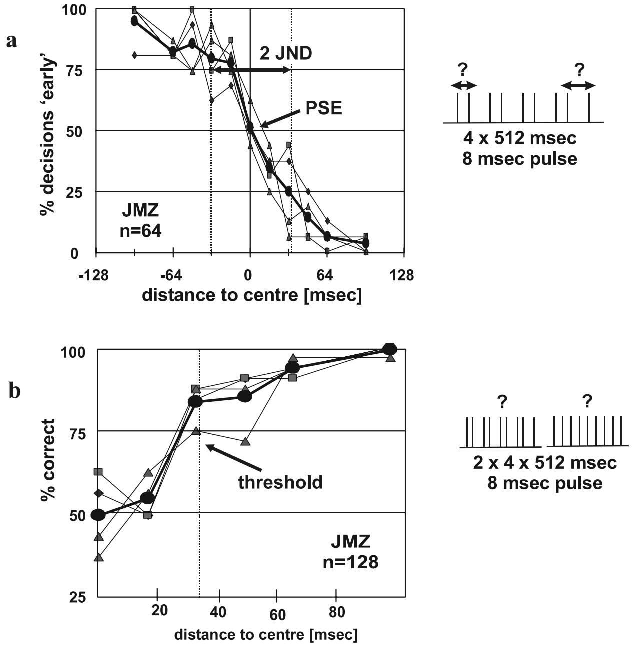

Fig. 5. (a) Psychometric function for subject JMZ and the repetitive temporal bisection task (sketched as inset). The frequency of decisions that, in a pair of intervals marked by a set of three consecutive light pulses, the first one is shorter than the second one, is plotted as a function of the presentation time of the second pulse in each triplet (given as distance from the centre). Small symbols and thin lines show the result from a single block; the black dots and thick lines show the average data from four blocks, which were used to estimate the PSE and JND. (b) Psychometric function for subject JMZ and a 2IFC detection (sketched as inset). The frequency of correct decisions whether the first or second of a pair of stimuli was composed of unequal duration intervals is plotted as a function of difference in duration. Small symbols and thin lines show the result from a single block; the black dots and thick lines show the average data from four blocks, which were used to estimate the threshold (broken line).

can be directly compared to the psychometric curves plotted for a single stimulus cycle in Fig. 1a. It is obvious from such a comparison that there is certainly no improvement through stimulus repetition, but rather some drop in performance as shown by a much shallower average psychometric curve. Correspondingly, the temporal position threshold (JND) increased from 25.3 to 33.0 ms (from 34.1 to 45.2 ms for the second subject, SCZ). This means that there is no extra benefit from multiple stimulus cycles in our experiment.

A second concern is the nature of the task employed in our main experiment. It might be that a simpler task, such as the detection of unequal intervals, would reveal higher temporal acuity than the discrimination of two intervals as is required to detect temporal position. Such an improvement would be very interesting because it would suggest an encoding of the existence of temporal phase shifts separate from that of the sign of such shifts. In the spatial domain, such separate encoding appears not to exist, at least for attention-demanding tasks such as those used here: instead, a single bipolar mechanism detects both departures from colinearity and the direction of a Vernier offset (Harris & Fahle, 1995). However, there might be an important difference in this respect in the visual coding of space and time. Perhaps also, in the apparently simpler task of detecting an irregularity of temporal intervals, the confusion about the order of intervals of different length, which was created by stimulus repetition in our first comparison experiment, would be replaced by some temporal integration over several regular or irregular stimulus cycles. Therefore we carried out a second comparison experiment that employed a simplified detection task.

In this test, subjects had to decide in a 2IFC paradigm which of two consecutive stimuli, presented in random order, contained a irregular sequence of intervals (SLSLSLSL or LSLSLSLS), as compared to a regular set of eight intervals that all had an identical intermediate duration of $M = (\mathrm{S} + \mathrm{L})/2$. All other stimulus parameters and aspects of the procedure, as well as the subjects, were identical to those of the previous comparison experiment. The task of detecting the irregular sequence seemed to the observers to require much less effort than that of discriminating the order of different length intervals (as in the main experiment and first comparison experiment), but their performance was no better. The psychometric curves, pooled for positive (LS) and negative (SL) differences between first and second interval durations are shown in Fig. 5b for the same subject and four repetitions of this experiment. The average psychometric curve resembles that shown for the discrimination task in Fig. 5a, reaching a threshold of 75% correct decisions in the same region as the temporal shifts leading to reliable responses in the first comparison experiment. The temporal position thresholds (JNDs) estimated with this task, 35.7 ms for JMZ and 32.0 ms for SCZ, did not show any substantial improvement relative to the values estimated from the single cycle temporal bisection task (25.3 ms and 34.1 ms, respectively). This finding further supports the view that the temporal position thresholds reported here reflect the limits of temporal encoding for the present stimulus configuration rather than the difficulty of that particular version of the task.

## 4. Discussion

In summary, our comparison of the three performance measures for different temporal judgements indicates that we failed to find in the human visual system hyperacuity for the temporal position of events. The precision with which our observers could estimate the temporal phase of a light signal relative to others does not go beyond the temporal resolution limit for separating two brief flashes or detecting the temporal modulation of a train of flashes. It should be remembered that the temporal bisection task is fundamentally different from a temporal two-point resolution task, because it employs temporal phase as the experimental variable, and thus provides a direct estimate of temporal position acuity. This lack of hyperacuity in temporal vision for stimulus configurations, which produce robust effects for spatial vision (Klein & Levi, 1985), raises a number of questions about the purpose and the mechanisms of localisation in space and time.

### *4.1. The neural basis of temporal coding*

First of all, one might argue that the purely formal analogy between space and time is just not helpful in understanding visual processing, because mechanisms which produce hyperacuity from spatial light distributions (Morgan, 1991) have no natural counterpart in the temporal domain. However, we believe that it would be not difficult to construct the temporal equivalent of a localisation mechanism that goes beyond the resolution limit for periodic modulations. The resolution limit is described by the sampling theorem (Bracewell, 1986) that requires at least two sampling points to cover each cycle of the periodic stimulus, the so-called Nyquist limit. This limit is set by the minimum receptor separation (distance) in the spatial domain and by the minimum spike separation (interval) in the temporal domain. Spatial localisation going beyond the Nyquist limit is achieved by a cortical mechanism comparing the relative activity levels of neighbouring input elements that receive optically blurred intensity profiles (Morgan, 1991; Morgan & Aiba, 1985).

All the necessary ingredients for such a mechanism in fact do exist in the temporal domain. Incoming light modulations are temporally blurred through the inherent low-pass properties of the receptors and primary sensory neurons (Van de Grind, Grüsser, & Lunkenheimer, 1973), leading to typical activity rates around 100 spikes per second and cutoff frequencies of approximately 50 Hz in the early visual system (Hamilton, Albrecht, & Geisler, 1989; Hammett & Smith, 1992; Metha & Mullen, 1996). Given these properties, it would be easy to design an operator that compared the phase angle of the pattern of averaged temporal activity in several neurons to a standard. So it is not surprising that such a mechanism, which exploits the phase relationships of averaged temporal sequences, is actually used by barn owls for auditory spatial localisation, requiring extreme precision in detecting inter-aural time differences (Knudsen, 1982; Wagner, Takahashi, & Konishi, 1987). Processes similar to that studied neurophysiologically in the barn owl could be the basis of high performance of humans in judging rhythms in the auditory system. Thus the neural machinery for temporal hyperacuity could be available in principle to the human visual system in contexts unrelated to spatial localisation, to other biological systems, and perhaps even to other sensory systems in humans, achieving temporal position acuity far beyond that observed in our present experiments.

### *4.2. The limits of temporal coding*

Second, it could be argued that there are biological limits to the temporal precision with which signals can be encoded in the human visual system or that the full potential of the neuronal processing machinery is not exploited in visual tasks. The experimental evidence, however, speaks against this interpretation. As noted in the Introduction, a range of studies supports the idea that the lower temporal threshold for detecting displacements in space and time can be as low as a few milliseconds (Sweet, 1953; Wehrhahn & Rapf, 1992; Westheimer & McKee, 1977). Similar thresholds are reached for change detection in non-spatial domains, like colour in the visual system or pitch in the auditory system (Yund & Efron, 1974). Even higher precision, approaching 1–2 ms, is achieved in the spatio-temporal interpolation paradigm in which a minute asynchrony between two lines is perceived as apparent Vernier displacement (Burr, 1979; Fahle & Poggio, 1981; Morgan & Fahle, 2000; Morgan & Watt, 1982; Morgan & Watt, 1983). Thus, when stimulus properties like location or colour are changed as markers for temporal order, human observers reach a temporal precision of a few milliseconds. This limit clearly goes beyond that observed in our experiments with a purely temporal task, in the absence of spatial offsets or other additional cues.

In another intriguing example, temporal factors have been studied in the context of feature binding through synchronisation as a stimulus feature that is used for figure–ground segregation (e.g., Kandil & Fahle, 2001; Usher & Donnelly, 1998). This task, however, involves processing at a higher level than the pure detection of temporal position which we were concerned within our experiments. If temporal precision used for segregation goes beyond conventional resolution limits, it could mean that highly acute temporal information is indeed available in the visual system. However, such information would only be retrieved for tasks such as figure-ground segregation, whereas the processes underlying temporal bisection would have no access to it. In contradiction to this view, when potential stimulus artefacts such as apparent motion are minimised, the best temporal precision under most conditions of figure–ground segregation is comparable to the limits defined by the flicker fusion frequency (Usher & Donnelly, 1998). On the other hand, at low oscillation frequencies a minimum phase shift between figure and ground of 15 ms can be reached that is sufficient for segregation. This value is better than the minimum half-period of 22 ms for a counterphase flickered figure–ground stimulus that leads to segregation, but still in the range of the minimum half-period of critical flicker fusion of 19 ms (see Kandil & Fahle, 2001). Because such lower limits are not reached at higher oscillation frequencies, this observation may reflect onset transients that might be used for comparison of neighbouring locations in the segregation task (see also Beaudot, 2002), suggesting some kind of second-order mechanism comparing temporal changes of higher stimulus features across space. Again, our single-location task turns out to be critical for assessing pure temporal acuity, because it excludes mechanisms that rely on phase shifts between stimulus changes at two separate locations.

### *4.3. The ecological importance of temporal coding*

If the neural machinery needed for millisecond precision is present in the human cortex, and if such precision is actually reached in different perceptual contexts, why does the present experiment fail to demonstrate a purely temporal hyperacuity? The clue may be given by asking what the ecological significance of any potential capacity for temporal hyperacuity might be. In vision, the precise temporal position of an event is not really significant to a living organism, whereas it might be in hearing (e.g., in speech). This is quite unlike detecting the spatial location of an object (such as spotting potential food, a predator or an obstacle), which obviously is of highest importance to any animal. Temporal position, or temporal order, does become highly relevant when things change (e.g., in colour or location) and the direction of change can be used to predict the future by extrapolation. It is obvious that fast and highly accurate temporal encoding of changes improves the reliability of such predictions. And it is perhaps most obvious that enhanced sensitivity for displacements across space and time requires high temporal and spatial precision, and that this sensitivity is critical for a number of behaviours that are crucial for survival. In this view, the extraordinary temporal precision shown in apparent motion tasks could be interpreted simply as a by-product of a highly sensitive spatio-temporal filtering mechanism which allows the visual system to deal with intricate, complex and rapidly changing patterns of motion signals in a fast, flexible, and robust manner. A challenge for future research will be to uncover instances of temporal precision in the hyperacuity range which oppose this rather general view of the tight connection between space and time in human visual processing.

## Acknowledgements

This project is difficult to locate temporally, because it started a long time ago and was diffused in space and time by substantial academic movements. So some of the work was done in the distant past, at different institutions around the world such as University College London or Australian National University, Canberra. We want to thank the various host institutions, patient subjects, and critical colleagues for contributing in manifold ways to this study. We further appreciate some insightful and constructive comments from our anonymous referees. Our special thanks go to Walter Junger for demonstrating his skills in designing the excessively simple experimental equipment.

## References

Beaudot, W. H. A. (2002). Role of onset asynchrony in contour integration. *Vision Research, 42*, 1–9.

Both, R., & Campenhausen, C. V. (1978). Sensitivity of a sensory process to short time delays. *Biological Cybernetics, 30*, 63–74.

Bracewell, R. N. (1986). *The Fourier transform and its applications*. Singapore: McGraw Hill.

Burr, D. C. (1979). Acuity for apparent Vernier offset. *Vision Research, 19*, 835–837.

Exner, S. (1875). Experimentelle Untersuchungen Der Einfachsten Psychischen Processe. III. Der Persönlichen Gleichung Zweiter Theil. *Pflugers Archiv*, 403–432.

Fahle, M., & Poggio, T. (1981). Visual hyperacuity: spatiotemporal interpolation in human vision. *Proceedings of the Royal Society London B, 213*, 451–477.

Finney, D. J. (1962). *Probit analysis*. Cambridge: Cambridge University Press.

Hamilton, D. B., Albrecht, D. G., & Geisler, W. S. (1989). Visual cortical receptive fields in monkey and cat: spatial and temporal phase transfer function. *Vision Research, 29*, 1285–1308.

Hammett, S. T., & Smith, A. T. (1992). Two temporal channels or three? A re-evaluation. *Vision Research, 32*, 285–291.

Harris, J. P., & Fahle, M. (1995). The detection and discrimination of spatial offsets. *Vision Research, 35*, 51–58.

Kandil, F. I., & Fahle, M. (2001). Purely temporal figure–ground segregation. *European Journal of Neuroscience, 13*, 2004–2008.

Kelly, D. H. (1972). Flicker. In: D. Jameson & L. M. Hurvich (Eds.), *Handbook of sensory physiology VII/4 visual psychophysics* (pp. 273–302). New York: Springer Verlag.

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validitation. *Journal of the Optical Society of America A, 2*, 1170–1190.

Knudsen, E. I. (1982). Auditory and visual maps of space in the optic tectum of the owl. *The Journal of Neuroscience, 2*, 1177–1194.

Metha, A. B., & Mullen, K. T. (1996). Temporal mechanisms underlying flicker detection and identification for red–green and achromatic stimuli. *Journal of the Optical Society of America A, 13*, 1969–1980.

Morgan, M. J. (1991). Hyperacuity. In: D. Regan (Ed.), *Vision and visual dysfunction 10. Spatial vision* (pp. 87–113). Houndsmill: Macmillan Press.

Morgan, M. J., & Aiba, T. S. (1985). Vernier acuity predicted from changes in the light distribution of the retinal image. *Spatial Vision, 1*, 151–161.

Morgan, M. J., & Fahle, M. (2000). Motion-stereo mechanisms sensitive to inter-ocular phase. *Vision Research, 40*, 1667–1675.

Morgan, M. J., & Watt, R. J. (1982). Hyperacuity for luminance phase angle in the human visual system. *Vision Research, 22*, 863–866.

Morgan, M. J., & Watt, R. J. (1983). On the failure of spatiotemporal interpolation: a filtering model. *Vision Research, 23*, 997–1004.

Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Perception, 24*, 1177–1187.

Sweet, A. L. (1953). Temporal discrimination by the human eye. *American Journal of Psychology, 66*, 185–198.

Tritsch, M. F. (1992). Fourier analysis of the stimuli for pattern-induced flicker colors. *Vision Research, 32*, 1461–1470.

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature, 394*, 179–182.

Van de Grind, W. A., Grüsser, O.-J., & Lunkenheimer, H. (1973). Temporal transfer properties of the afferent visual system. In: R. Jung (Ed.), *Handbook of sensory physiology VII/3. Central processing of visual information A* (pp. 431–573). Berlin: Springer.

Wagner, H., Takahashi, T., & Konishi, M. (1987). Representation of interaural time difference in the central nucleus of the Barn Owl's colliculus. *The Journal of Neuroscience, 7*, 3105–3116.

Wearden, J. H., & Bray, S. (2001). Scalar timing without reference memory? Episodic temporal generalization and bisection in humans. *The Quarterly Journal of Experimental Psychology, 54 B*, 289–309.

Wehrhahn, C., & Rapf, D. (1992). ON- and OFF-pathways form separate neural substrates for motion perception: psychophysical evidence. *The Journal of Neuroscience, 12*, 2247–2250.

Westheimer, G. (1984). Spatial vision. *Annual Reviews of Psychology, 35*, 201–226.

Westheimer, G., & McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research, 17*, 887–892.

Yund, E. W., & Efron, R. (1974). Dichoptic and dichotic micropattern discrimination. *Perception and Psychophysics, 15*, 383–390.